DATA ANALYTICS MINI-PROJECT SYNOPSIS

**Project Title: Car Evaluation System**

**Team Members:** 1. Vaibhav Gole (402062)

2.Rishikesh Kakad (402075)

3. Yoganand Kanhed (402077)

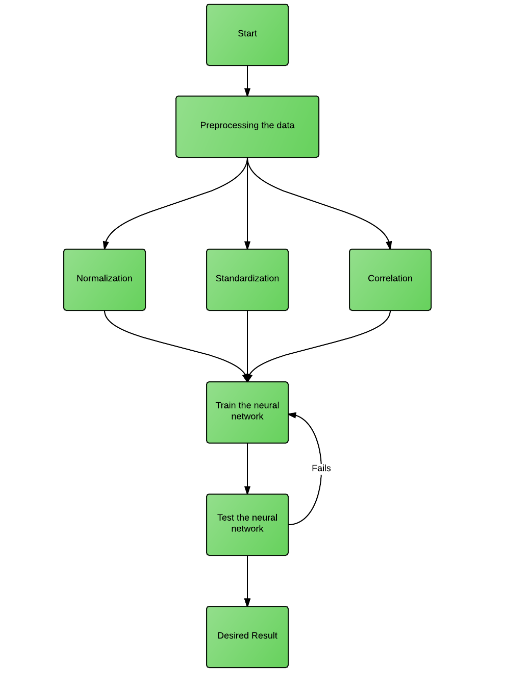
**Project Domain:** Data Analytics ,Neural network

**Project Description:**

The main reason in building the decision making model is to predict the price of car. Nowadays predicting the price of a car is a daunting task. A car contains lot of assemblies and equipments which are of various values and from different generations. In addition to these in some cars remodeling will be done which will complicate the process. To solve this, we designed a system which will predict the price of the car based on its attributes.

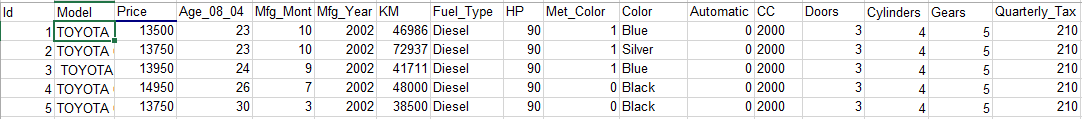
**Input**: Different Attribute of Car Dataset.

**Output**: The Help of Neural Network predicting car price.

**Flowchart:**

**Input Dataset**

It consists of all the characteristics of the car necessary for the evaluation of the car which are considered as the attributes in the dataset. In this dataset, there are thirty-eight attributes like Model, Price, Age, No. of Cylinders, Airbags etc., are shown in the figures. These are divided into categorical-type, numerical-type and constant-Type



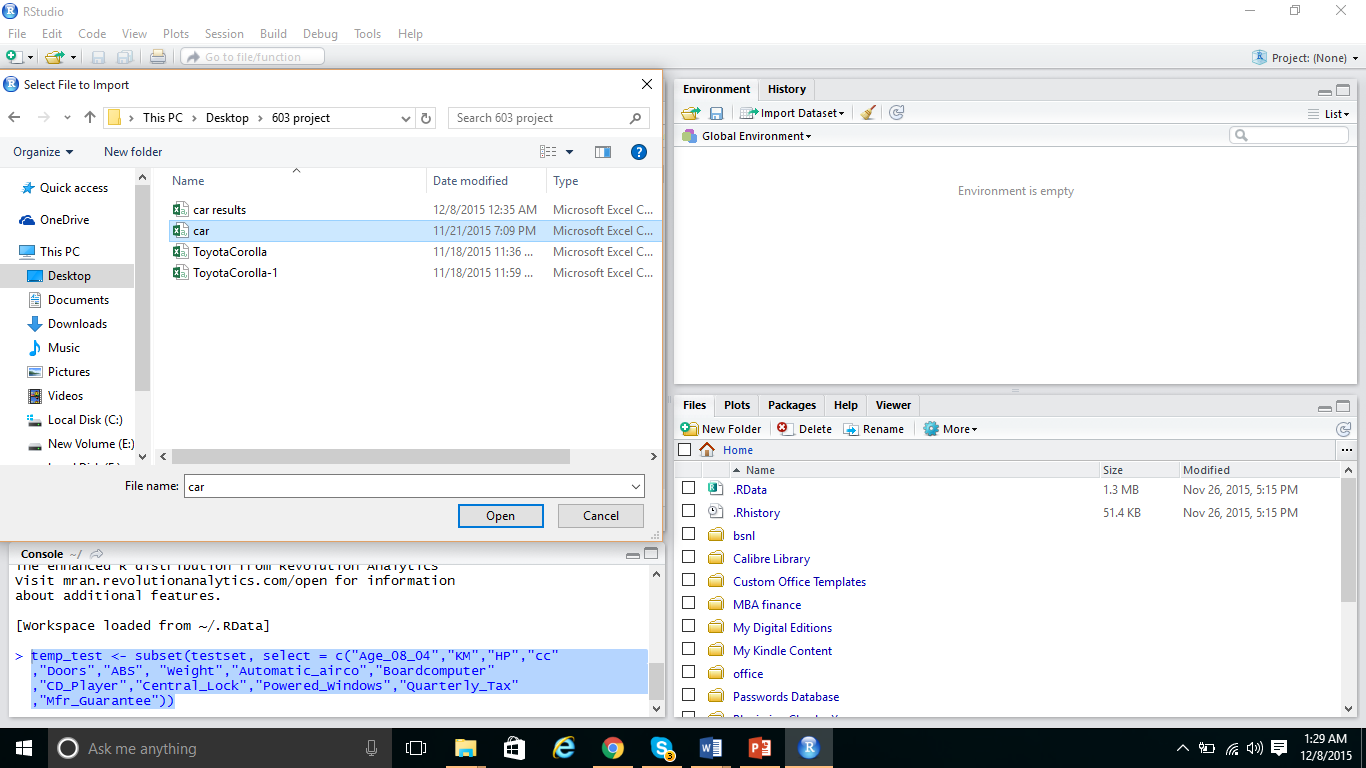
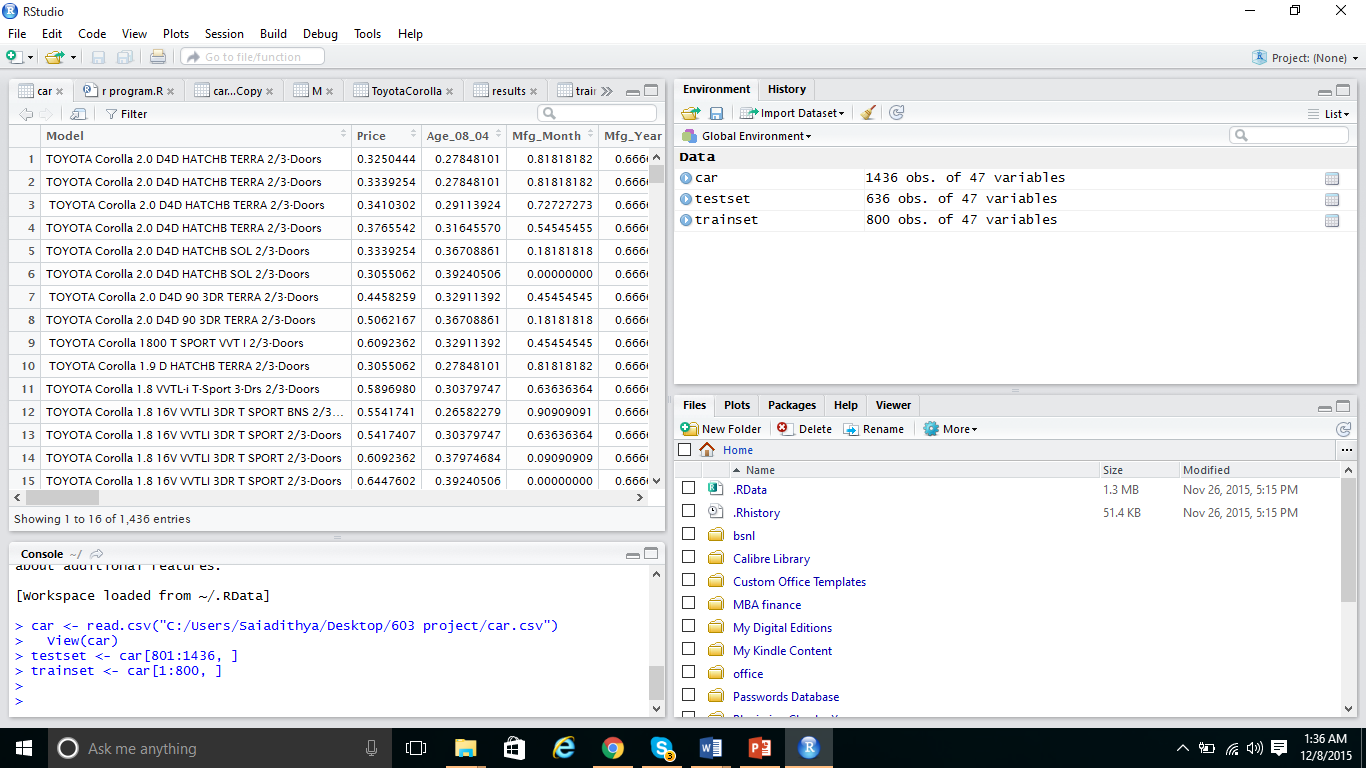
attributes based on their characteristics.

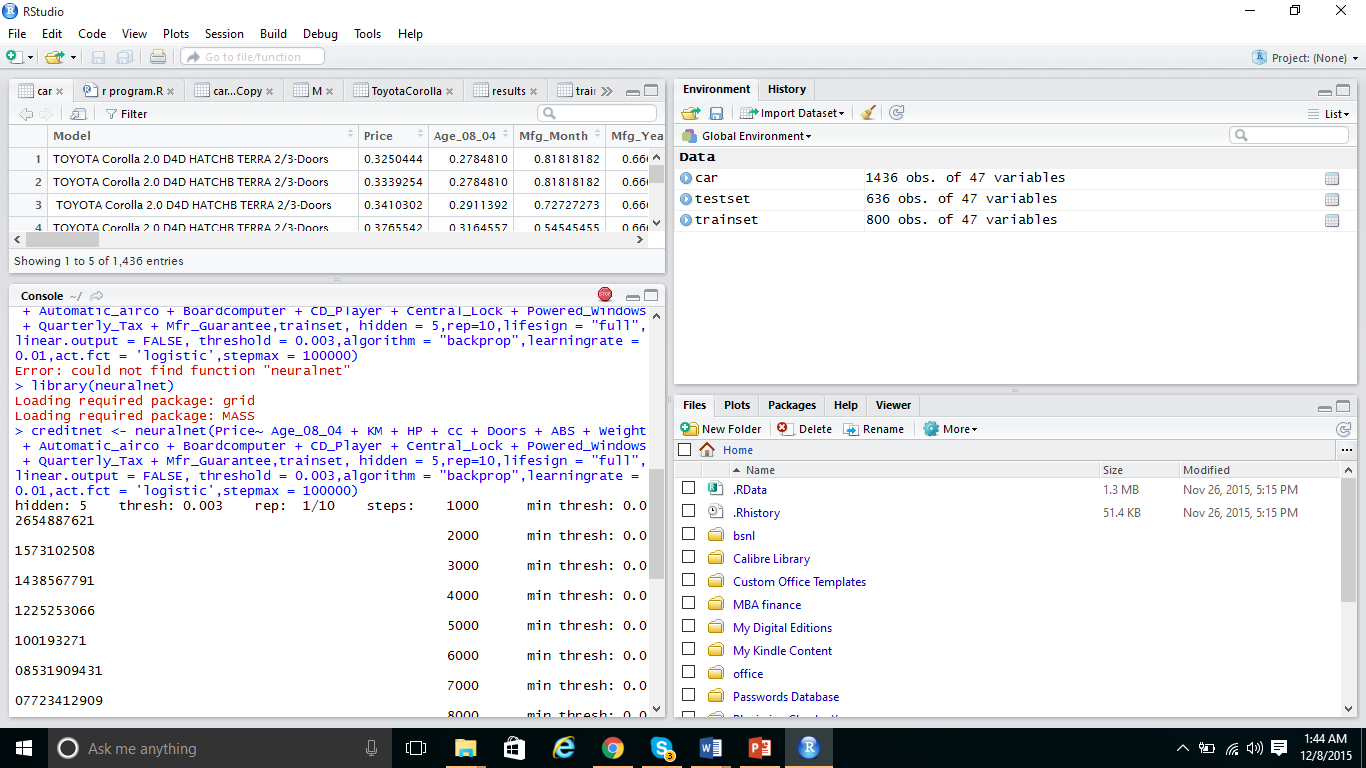
**Pre-Processing:**

Pre-processing of the data takes place in two stages:

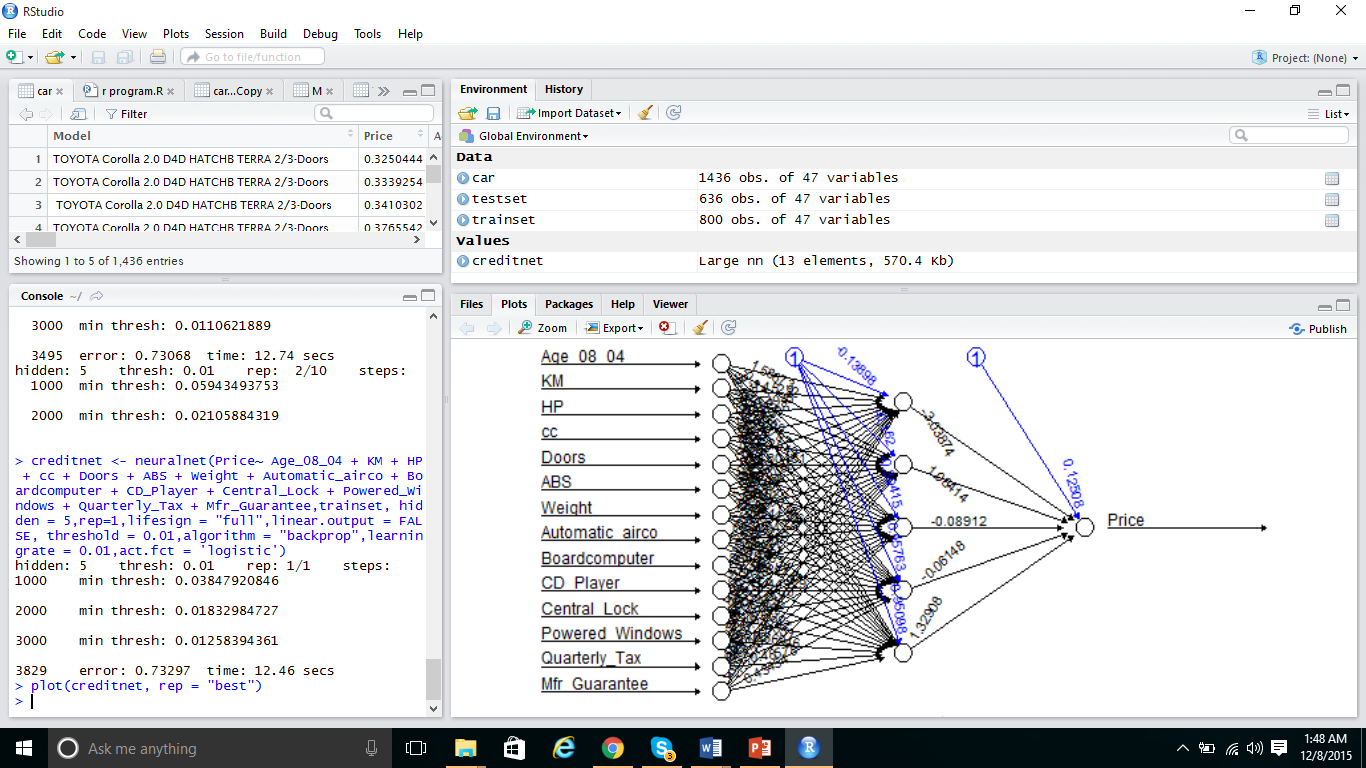
* Standardization
* Normalization

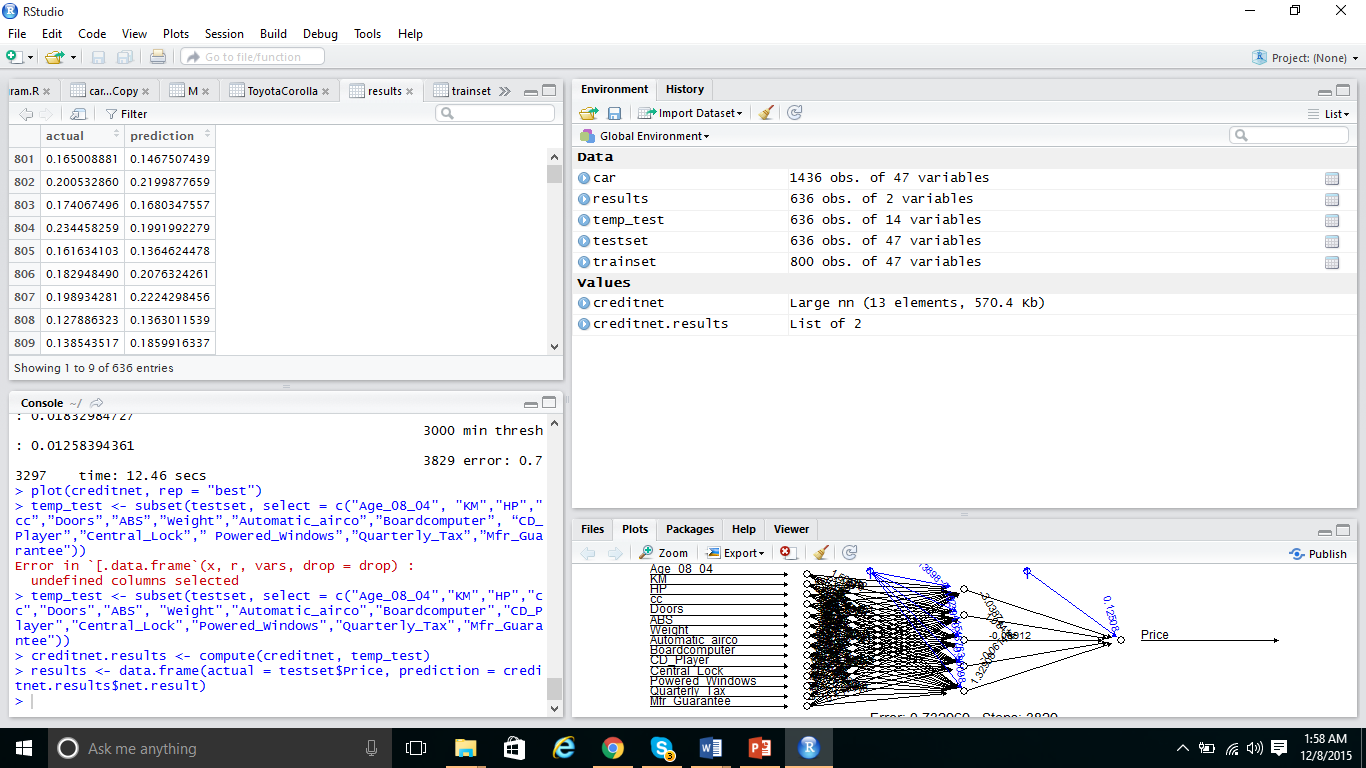
**Implementation Steps:**

* Import the input dataset for the experiment.
* Then separate the dataset into 2 sections. One is for training and testing.
* Construct the Neural network with the initial parameter values found during the training using the Neural net package formula.



* Plot the Neural network graph using Plot formula in R studio.



* It is then run on Temporary dataset (Temp test) which is created copying the contents of the Test dataset into it. The purpose of doing this is to separate other unnecessary columns from the neural network computation
* Compare the Actual and predicted value of the neural network. Use the Compare code in R studio and View command to view the result.

**Neural Network Design Formula:**

Neural net package uses specific formula for designing a neural network.

creditnet <- neuralnet(Price~ Age\_08\_04 + KM + HP + cc + Doors + ABS + Weight + Automatic\_airco + Boardcomputer + CD\_Player + Central\_Lock + Powered\_Windows + Quarterly\_Tax + Mfr\_Guarantee,trainset, hidden = 5,rep=10,lifesign = "full",linear.output = FALSE, threshold = 0.009,algorithm = "backprop",learningrate = 0.01,act.fct = 'logistic',stepmax = 100000)

Where

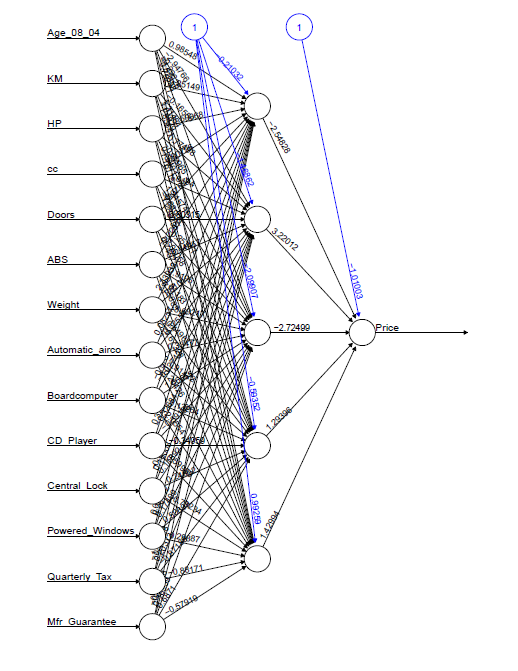
* Creditnet is the name of the Neural network to be designed
* Price, the output of the Neural network
* Age, KM,HO,CC etc. are the inputs to the network
* Hidden is a vector specifying the number of hidden layers and nodes

**R Code to compare Actual vs Predicted values:**

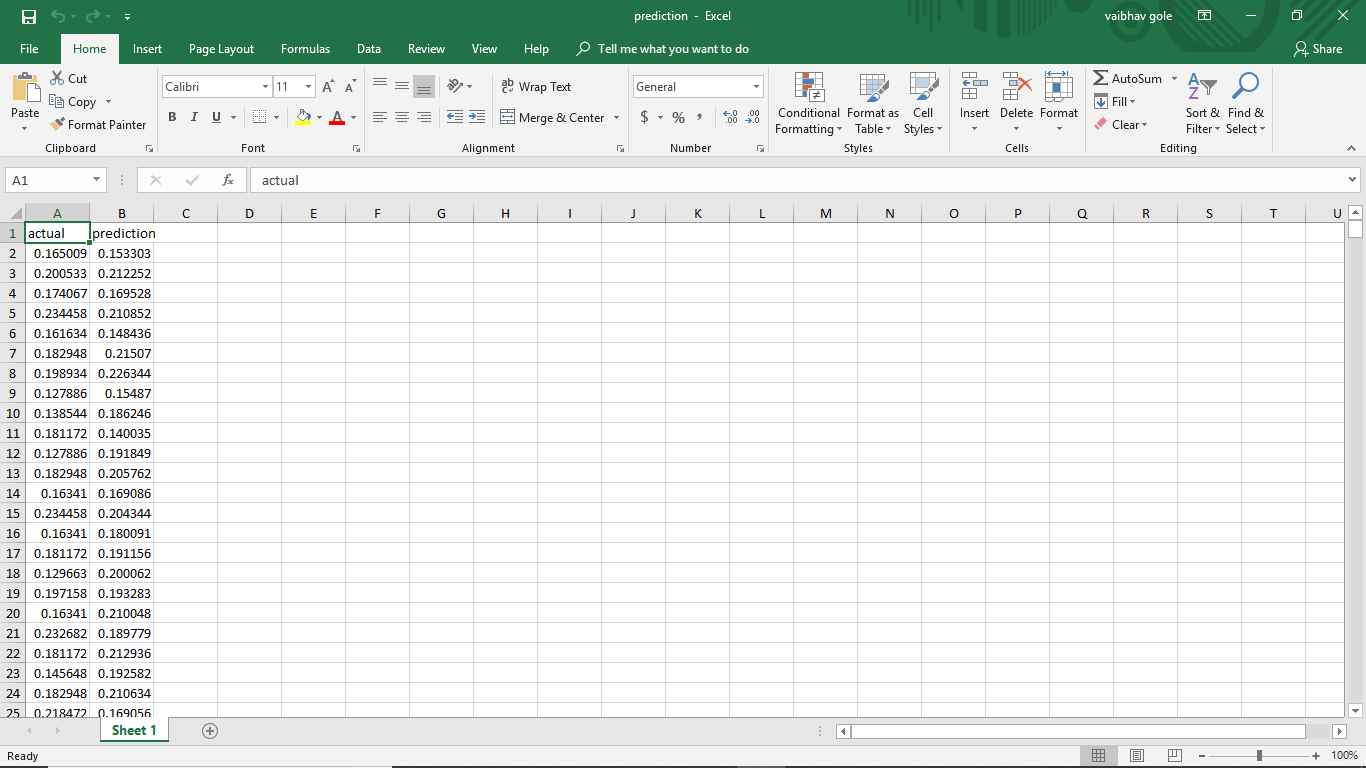
> results <- data.frame(actual = testset$Price, prediction = creditnet.results$net.result)

Code compares the Price attribute from the test dataset to the predicted output stored in the creditnet.results dataset.

**Output: Neural Network**

****

**Price Prediction:**

****

**Software Requirements:**

Software Tool:

Rapid Miner

R studio

**Hardware Requirements:**

Intel® Core™ i5-2430M CPU @ 2.40GHz 2.40GHz.

4.00GB RAM

**System Limitations:**

* Few of the predicted values varies from the actual by a huge margin which cannot be analyzed. This is because Neural network is
* Although constructing the neural network was easy, the real test lies in predicting the Network inputs and its parameters.
* Adding a new attribute in the dataset cannot be done as simple and it requires repeating
* the whole training and preprocessing steps.

**Conclusion:**

We have constructed a Neural network using R studio (Neural Net) package and for predicting the price of the car. From the project conducted, we understand that Neural network is a black box and to correct the problem we must carefully analyze each of its parameters and preprocess the data. We can close out the margin of error but we can never get a zero error in the neural network We improving the efficiency of the neural network by following some state of the art methods. Although for some car, the predicted price is huge margin from original price. Overall the system is very useful and quite efficient for yielding the car price for a range of attributes.

**Future Enhancements:**

System needs new methods to preprocess the data easily. Preprocessing is a daunting task and consumes more time. New methods to find the problem occurrence in neural networks should be introduced.